Le Centre canadien de recherche sur l'informatisation du travail

Canadian Workplace Automation Research Centre

Programme de recherche et développement du CCRIT en Traduction assistée par ordinateur

par

Pierre Isabelle
et
Elliott Macklovitch

Chercheurs

Centre canadien de recherche
sur l'informatisation du travail

Canada

Programme de recherche et développement du CCRIT en Traduction assistée par ordinateur

par

Pierre Isabelle
et
Elliott Macklovitch

Chercheurs

Centre canadien de recherche
sur l'informatisation du travail

Version anglaise disponible

CO28-1/21-1988F ISBN 0662-95-71-2

Il est généralement reconnu que la traduction entièrement automatique de bonne qualité de textes arbitrairement choisis (parfois appelée FAHQTUT) est un objectif inaccessible, tout au moins dans un proche avenir. Cette inaccessibilité s'explique par le fait que la traduction correcte de textes en langues naturelles exige systématiquement des connaissances extralinguistiques, c'est-à-dire des connaissances qui ne soient pas explicitement exprimées dans le seul texte. Si des progrès considérables ont été réalisés au cours des trente dernières années pour amener les ordinateurs à rendre le sens linguistique d'une phrase, nous en savons encore assez peu sur la façon de représenter, de mémoriser et d'appeler des quantités d'informations non linguistiques relatives au monde tel qu'on le connaît ou sur la façon d'amener une machine à "raisonner intelligemment" de manière à pouvoir traduire aussi bien qu'un traducteur, aussi novice soit-il.

Il s'ensuit donc que, pour permettre aux ordinateurs de jouer un rôle important dans la satisfaction de la demande de traduction mondiale sans cesse croissante, le but final de la FAHQTUT doit être réduit et des objectifs sous-optimaux, fixés pour la traduction assistée par ordinateur. Des compromis sont possibles dans les trois axes suivants :

. le degré d'automatisation,
. l'éventail de textes qu'un système peut traiter et
. la qualité des traductions qu'il produira.

De ces axes, le dernier est peut-être le moins intéressant : peu de gens sont prêts à accepter des traductions approximatives contenant des erreurs imprévisibles que seule une machine peut faire. Par ailleurs, on peut mettre au point des systèmes qui ne sont pas entièrement automatiques, soit en laissant au traducteur humain l'initiative de la traduction et en lui fournissant un jeu d'outils pour accomplir plus efficacement des tâches connexes telles que la recherche terminologique et le traitement de texte multilingue - on parle alors de traduction humaine assistée - ou en confiant l'initiative de

la traduction à l'ordinateur avec l'assistance humaine voulue pour résoudre les problèmes trop difficiles - on parle alors de traduction machine assistée par l'homme.

À ce jour, tous les systèmes de traduction automatisée à usage général ont nécessité un certain degré d'intervention humaine, soit avant, durant ou après le processus de traduction proprement dit. Dans la plupart des cas, le fardeau de cette interaction a été soit trop lourd pour que les traducteurs le supportent ou trop considérable pour que les systèmes soient rentables.

La troisième technique de sous-optimisation cherche à remédier à cette situation en sacrifiant l'universalité à la qualité. Les systèmes élaborés selon cette approche ne sont pas destinés à traduire des textes choisis arbitrairement. Ils sont plutôt conçus expressément pour des sous-langues restreintes, soit des sous-langues naturelles telles que les bulletins météorologiques et les rapports d'activité boursière ou des sous-ensembles définis de langues naturelles tels que certains types d'anglais simplifié qui imposent des restrictions linguistiques aux rédacteurs en vue de faciliter la compréhension des textes par l'homme ou par la machine. L'application la plus réussie de l'approche, dite par sous-langue, à la traduction automatisée est le fameux système MÉTÉO dont se sert le Bureau fédéral des traductions pour traduire tous les jours plus de 20 000 mots de prévisions météorologiques.

Le programme de l'équipe de R.-D. du CCRIT en traduction assistée par ordinateur est fondé sur une évaluation de l'état actuel de la technologie exposée ci-dessus. L'objectif de recherche à long terme du groupe est d'étudier des moyens de rendre les systèmes de traduction automatisée plus intelligents, c'est-à-dire moins littéralement liés à la formulation du texte d'entrée, en adaptant des techniques qui ont donné des résultats dans d'autres domaines de l'intelligence artificielle. Une attention particulière sera prêtée à la notion de règle de traduction dans la génération suivante

des systèmes de traduction automatisée et à la meilleure façon de représenter et d'acquérir le savoir contrastif interlinguistique.

À moyen terme, le groupe espère trouver une ou plusieurs sous-langues prometteuses pour lesquelles des systèmes spécialisés de traduction automatisée, fondés sur le modèle MÉTÉO, pourraient être mis au point. Et, enfin, un projet est actuellement en cours pour créer un poste de travail spécialisé en traduction, de conception ergonomique, avec un ensemble d'outils informatisés intégrés qui permettraient au traducteur d'aujourd'hui d'accomplir plus efficacement des tâches courantes liées à l'administration, à la recherche et au traitement de texte et d'être, par conséquent, plus productif. Si vous avez des questions, veuillez communiquer avec les auteurs.